Spring 2003

ESDCD news

Code 930 • http://esdcd.gsfc.nasa.gov

Earth and Space Data Computing Division

Earth Sciences Directorate, Goddard Space Flight Center

In This Issue

NCCS System Transitions

NCCS Retires T3E

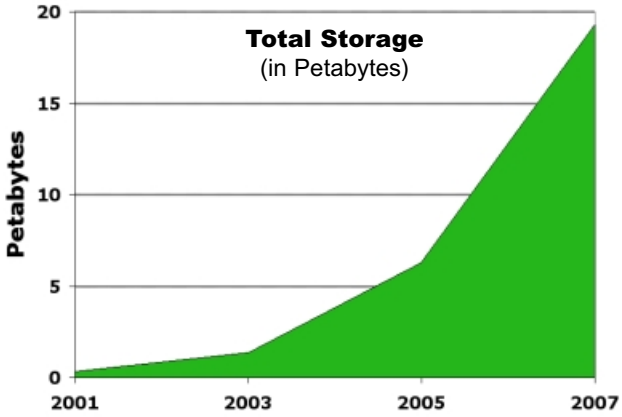
The NASA Center for Computational Sciences' (NCCS) Cray T3E, once the fifth most powerful computer in the world, was retired at the end of 2002 and deinstalled on March 5, 2003.

The acquisition of the Cray T3E in 1997 marked a major shift from vector supercomputing to large-scale parallel computing. By using commodity processors with a high speed interconnect, this system provided both significant capability and capacity. The T3E was a platform for research in high-performance computing methods. The system was used for the study of El Niño and related phenomena by the NASA Seasonal-to-Interannual Prediction Project (NSIPP).

A cabinet of the T3E system is loaded onto a truck. Image credit: Richard Glassbrook, AMTI

The 1,360-processor T3E with peak computing capability of 0.778 TeraFLOPS was replaced by a 1,392-processor HP/Compaq SC45 with peak computing capability of 3.2 TeraFLOPS (see Fall 2002 ESDCD News). The new system is shared by all the NCCS user groups.

NCCS Enters Petabyte Era with New Mass Storage System



NCCS observed and projected data storage total Petabytes (including risk mitigation duplicates).
Chart credit: Ellen Salmon

To meet the critically growing storage needs of the NCCS' Earth and space sciences research community, the NCCS has acquired a new high-end server and software that will expand the capabilities of the NCCS' mass storage system.

The volume of data from users' store and retrieve traffic on the current NCCS mass storage system has been doubling every 15–24 months and is pro-

jected to skyrocket to as much as 20 Petabytes (PB), or 20,480 Terabytes (TB), by the end of 2007.

This growth rate is being driven in part by researchers who have been able to increase the resolution of their models by using the NCCS' HP/Compaq AlphaServer SC45. Another growth factor is the impending arrival of workload and data from GSFC's Data Assimilation Office (DAO). The DAO's large scale, computationally intensive simulations of Earth processes are returning to the NCCS from NASA's Ames Research Center (ARC).

After an intensive year-long evaluation of vendors' technical solutions, the NCCS selected the new Sun Fire 15000 server and Sun StorEdge Performance Suite and Utilization Suite (SAM-QFS) software. This system is considered the best long-term solution for the NCCS user community. The new equipment will replace the existing NCCS Sun E10000 system running Legato UniTree software.

According to Nancy Palm, who manages the NCCS, "We are embracing this new system because of the benefits that its expanded features will bring to our user community. In addition to increasing capacity of the mass storage system, the Sun solution will bring new expandability and functionality options to NCCS users, while increasing system accessibility and reliability."

The faster processors, improved input/output (I/O) bandwidth, and greater number of high-speed network interfaces on the new system will provide an infrastructure with a peak performance up to 3 times faster than the existing Sun E10000/UniTree system. As delivered, the new system is expected to accommodate peak I/O rates at least 5 times greater than those needed to sustain the projected 20 PB requirements in Fiscal Year (FY) 2007 (with nearly 8.5 PB added during FY 2007). To satisfy growth in requirements, the production system can expand to accommodate up to 32 additional Gigabit (Gb) Ethernet or Fibre Channel ports, an increase of 80 percent over the delivered configuration.

A new 9.3-TB DataDirect Networks S2A 8000 RAID disk array will provide the initial disk storage for the new system. The new system's disk space will expand to 18.6 TB when a second 9.3-TB S2A 8000 is added after all UniTree data have been copied into SAM-QFS.

Although the new server will appear as one system to NCCS users, it will actually be a clustered configuration of two productions domains, each with access to all of the system's disk and tape storage. SAM-QFS, configured with Veritas Cluster Server software, will allow access to all data during system maintenance and outages. The server's fault-isolation and hardware redundancy features will minimize entire-server outages.

High performance and scalability are additional benefits of the new server. Each of the two production domains will have eight 1,050-MHz UltraSPARC III CPUs, 16 Gigabyte (GB) of memory, three Jumbo Frame-capable Gb Ethernet network interface cards, and twenty-four 2-Gigabit (Gb) Fibre Channel ports, with additional room for expansion.

The NCCS will more than triple the potential capacity of its nine StorageTek Powderhorn robotic tape silos to over 9 PB of data with the addition of StorageTek T9940B Fibre Channel tape drives. The new drives will write 200 GB on a tape cartridge at a 30-Megabyte-per-second (MB/sec) transfer rate, while the outgoing T9940A drives write 60 GB per tape with a 10-MB/sec transfer rate.

The NCCS has developed a comprehensive three-phase data migration plan to transfer the UniTree HSM data to SAM-QFS transparently. During all three phases, the NCCS will work closely with its user community to ensure that data are transferred smoothly and with minimal disruption.

During the first phase in early Summer 2003, both new SAM-QFS data and existing UniTree data will be accessed via the SAM-QFS system. In the second phase, to be completed by September 30, 2003, existing UniTree data will be copied into the SAM-QFS system. During the third phase, in FY 2004, all DAO Data Migration Facility data will be migrated into SAM-QFS.

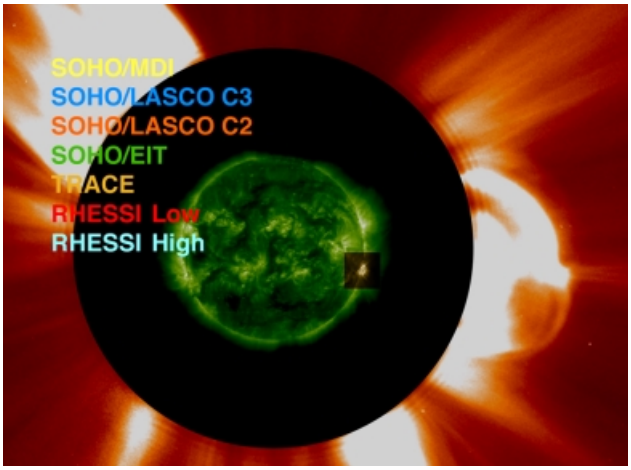
http://esdcd.gsfc.nasa.gov/SCB

SVS

Multi-mission Solar Movie

The Scientific Visualization Studio (SVS) produced a visualization that is one of the first to combine data from a fleet of Sun-observing satellites focused on a single event. "Multi-mission Solar Movie" is a data-driven sequence of AR9906, a solar flare that occurred April 21, 2002, on the Sun's western limb. The movie was composited using data sets from seven instruments onboard three satellites.

"Satellites working together in this manner may eventually help reveal how solar cycles and variations impact Earth's climate," says William Bridgman of GST, an SVS data visualizer with a background in gamma ray astrophysics.



An expanding bubble of hot plasma expands into the SOHO/LASCO camera 2 (C2) field of view. Image credit: SVS

Bridgman has been producing visualizations for the Sun-Earth Connection (SEC) program for over 2 years, with funding from the SEC Division of NASA's Office of Space Science. The SEC Division seeks to understand the Sun, the heliosphere, and planetary environments as a single connected system.

The visualization is a seven-instrument composite, one of the first to include x-ray data from the Ramaty High Energy Solar Spectroscopic Imager (RHESSI). RHESSI was launched in February 2002 to explore the basic physics of particle acceleration and explosive energy release in solar flares.

The sequence includes data from instruments on-board the Solar and Heliospheric Observatory (SOHO) satellite, a cooperative effort between the European Space Agency and NASA to explore the Sun. Included are visible wavelength data from the Michelson Doppler Imager (MDI), outer solar corona data in visible wavelengths from the Large Angle and Spectrometric COronagraph (LASCO), and ultraviolet (UV) data from the Extreme ultraviolet Imaging Telescope (EIT).

Finally, UV data from the Transition Region and Coronal Explorer (TRACE) satellite are included. TRACE is part of NASA's Small Explorer (SMEX) mission to image the solar corona and transition region at high angular and temporal resolution.

The 30-second, 880-frame sequence starts prior to the solar flare and then shows the spectacular coronal mass ejection (CME) billowing out to the side. The movie transitions from one set of satellite data to another, with a full composite of all the data towards the end of the sequence.

Bridgman processed over 86 GB of high-resolution, nearly raw, unsynchronized data obtained from GSFC's online Solar Data Analysis Center (SDAC) and from instrument teams. The data were captured from the different instruments at rates ranging from every 6 hours to every 12 seconds and contained artifacts unique to each instrument.

To prepare the data for rendering, Bridgman removed the artifacts by experimenting with interpolation routines and other algorithms. He then automated synchronization of the data using Python, an open source, object-oriented scripting language.

Contributors included Brian Dennis and Joseph Gurman of GSFC's Laboratory for Astronomy and Solar Physics; Bernhard Fleck of SOHO; Alexander Kosovichev and Todd Hoeksema of Stanford University; Steele Hill and Peter Gallagher of L-3 Communications Analytics Corporation; and Nathan Rich of Interferometrics Inc.

GSFC's Public Affairs Office recently released this visualization on its annual SEC resource reel for use by the news media.
http://svs.gsfc.nasa.gov/search/AnimationSeries/Multi-MissionSolarMovie.html

SVS Open House: Expanded Facilities

The SVS held an open house on February 13, 2003, to announce its expanded location and to commemorate 15 years of providing a wide range of visualization expertise to NASA. The multimedia event showcased SVS projects from past and present. Visitors were treated to tours of the facilities, large theater displays, virtual reality and cluster computing demonstrations, full-color SVS posters exhibited across the building, and live video streaming demonstrations.

"We are really proud to host this fun and exciting event and to welcome back friends of the studio, as well as to let the NASA community know about our expanded facilities," says Horace Mitchell, Manager of the SVS.

Representatives from GSFC's Public Affairs Office (PAO) were on hand to help describe the types of visualizations produced at the SVS. Working closely with the studio, PAO requests and releases products for use by a wide range of media to help explain NASA projects and missions. SVS products have been integrated into scenes of popular movies and numerous science documentaries for broadcast by the mass media, such as the SCI FI and Discovery channels. They are also often featured in publications such as National Geographic, Science News, and Science.

The new SVS facility featured its 50" High Definition Television projection screen. The screen enables the SVS and its clients to preview the studio's high-definition animations in full resolution. The SVS also uses the screen to demonstrate its products to groups of up to 40 people.

The SVS's six-screen visualization cluster. Image credit: SVS

Visitors toured the SVS where scientific visualization experts gave informal talks about how SVS visualizations are used by researchers and educators to help describe scientific results in publications and outreach efforts. The SVS displayed data-driven visualizations, while Goddard TV's Conceptual Image Lab exhibited animations of NASA science and missions, including dramatic concepts for NASA's SEC program.

Visitors watched demonstrations on leading-edge visualization equipment. Landsat images comparing Mt. Kilimanjaro snowcap levels were shown on a six-screen visualization cluster of six PCs designed by SVS personnel. Also located in this area is the SVS's Virtual Reality workbench, which is used to help maintain the Digital Earth (DE) prototype. According to Mitchell, the SVS is using the workbench to port the DE software to a PC as part of an effort funded by NASA's ARC 2003 Learning Technologies Project.

Finally, visitors saw live video streaming demonstrations from the Goddard TV Studio. The SVS regularly produces background visualizations for Goddard TV, where live interviews with scientists and other experts are streamed directly to broadcast media. This studio is conveniently located in the same building as the SVS. **http://svs.gsfc.nasa.gov**

High Performance Networking Forefronts

SAN-Over-IP

A GSFC-wide pilot effort is underway to explore advanced storage architectures to help retain and distribute NASA's large and growing data holdings. Initially encouraged by Milton Halem during his tenure as CIO of GSFC, and led by Benjamin Kobler of GSFC's Science Data Systems Branch, the effort focuses on testing and developing Storage Area Networks (SAN). The ESDCD's High End Computer Network (HECN) Team, led by Patrick Gary, is extending this effort to use SANs over high-speed Internet Protocol (IP) networks.

SAN-over-IP can bring many benefits to users. It can enable access to isolated storage infrastructures, or SANs, connected to IP networks. SAN-over-IP can allow data at one location to be shared globally among multiple SAN-connected clients on different platforms from within GSFC, as well as from a broader, more geographically dispersed user base. Furthermore, with the growing use of Gb/sec fiber optic networks, data accesses of SAN-over-IP can approach the speed of local 1-2 Gb/sec Fiber Channel-based SANs. Other benefits of SAN-over-IP may include cost savings associated with consolidated storage, greater utility of centralized data, and better data protection through replication.

Kobler's team installed a local SAN with Fibre Channel switches across three buildings at GSFC. Kobler demonstrated that client computers attached via Fibre Channel host bus adapters could access disks attached to the SAN in any of the three buildings. Data reading and writing averaged approximately 50 MB/sec.

However, SANs created with Fibre Channel can be extended to only 10 kilometers (km). The HECN team designed and implemented a cost-effective pilot configuration that extended GSFC's local SAN beyond 10 km to include SAN-over-IP users outside of GSFC. HECN achieved this extension to the University of Maryland, College Park, by using Internet Small Computer Serial Interface (iSCSI) protocols and software device drivers in the public domain. These tools were used together with a Cisco 5420 storage router attached directly both to GSFC's local Fibre Channel SAN and to GSFC's Gigabit Ethernet (GE)-based Scientific and Engineering Network (SEN).

The extended SAN-over-IP configuration includes PCs, Macintoshes, and Suns, running iSCSI over their normal GE network interface cards (NIC).

iSCSI-based data reading and writing test results averaged approximately 25 MB/sec, which still is very fast. Even though these results are about half as fast as the direct-Fibre-Channel-attached tests, these speeds were achieved across distances impossible for Fibre Channel-only SANs to span, and at significantly less cost.

The HECN team continues to develop GSFC's SAN-over-IP with the goal of extending to more distant sites. A white paper detailing GSFC's SAN and SAN-over-IP pilot configurations and test results was presented at the 20th IEEE Symposium on Mass Storage Systems, held in cooperation with the 11th NASA Goddard Conference on Mass Storage Systems and Technologies in April. The paper, "Considerations and Performance Evaluations of Shared Storage Area Networks at NASA Goddard Space Flight Center," was prepared by Hoot Thompson of Patuxent Technology Partners; Curt Tilmes of GSC; Robert Cavey of ICS; William Fink of ESDCD's Science Communications Technology Branch; Paul Lang of ADNET; and Kobler.
http://storageconference.org/2003/papers/21-Thompson-Considerations.pdf

Outreach

Commendations: GLOBE Visualization Server Team

After over 8 years of successful development at GSFC, the Global Learning and Observations to Benefit the Environment (GLOBE) program has been awarded through a NASA cooperative agreement to the University Corporation for Atmospheric Research. This award was made to find partnerships that will expand GLOBE's resources and to introduce new methods that will further increase GLOBE's audience.

The GLOBE visualization team built and refined GLOBE's visualization server, one of the world's first automatic Web interfaces for on-demand generation of scientific data in real-time for educational purposes. Features include automated maps, graphs, and tables of environmental data collected by students from across the globe. In the past year, the server has responded to 3 million Web hits from 40 thousand different users. Over 100 nations have joined GLOBE, with data collected on every continent, including Antarctica.

In letters of recognition, Ronald Birk, Director of the Earth Science Applications Division of NASA's Office of Earth Science, commended GLOBE visualization team members for their outstanding performance in making GLOBE "...a truly remarkable program that has helped the education of over a million students worldwide...." Birk applauded their contributions to NASA's mission as the project "...continues to demonstrate how science and education can come together for the benefit of society."

Team members include project manager David Batchelor, Horace Mitchell, and Raul Garza-Robles of the ESDCD; Peter Jackson, Sally Stemwedel, and Phillip Keegstra of GST; Theresa Held of CSC; and Jeffrey Cieslak of SSAI.

Direct Readout Project

Patrick Coronado implemented Direct Readout project technologies in the January 2003 JASON XIV expedition, "From Shore to Sea." Participants in this educational program explored the unique features and local coastal ecosystems of the California Channel Islands. The Direct Readout project enabled JASON for the first time to present near real-time Moderate Resolution Imaging Spectroradiometer (MODIS) data on chlorophyll, sea surface temperature, and other environmental factors. Coronado and the SVS also produced an educational video on remote sensing that was viewed by more than 1 million students in the U.S., Germany, and Mexico.
http://directreadout.gsfc.nasa.gov
http://www.jasonproject.org

MU-SPIN Initiatives

In order to increase students' access to science and technology, the Minority University-SPace Interdisciplinary Network (MU-SPIN) program regularly promotes and supports educational outreach activities, including the following March 2003 events:

- MU-SPIN supported the grand opening activities for the Center of Excellence in Remote Sensing Education and Research at Elizabeth City State University on March 4. This new remote-sensing laboratory will further scientists' research in coastal, ocean, and marine science and will prepare students for graduate studies and careers in these and related fields of study.

- MU-SPIN staff coordinated NASA Science Week March 18–21 at Bennett College, a liberal arts college for women. Emphasizing key areas such as engineering, astrobiology, education, and forensic science, the event focused on space technology and exploration and how they impact the future of life on Earth. Aprille Ericsson, of GSFC's Flight Dynamics Analysis Branch, gave the keynote address. Professionals from across

the Agency shared their expertise with students through workshops, exhibits, and roundtable discussions.

- As part of the Network Resources and Training Sites Distinguished Lecture Series, MU-SPIN, Fayetteville State University, and Elizabeth City State University presented the lecture "Excellence Without Excuses" by Julian Earls, Deputy Director of NASA's Glenn Research Center. MU-SPIN staff provided a Web cast of the March 22 lecture, available at the MU-SPIN Web site.

MU-SPIN is also playing a lead role in the development of an integrated distance-learning program for NASA. Building on the "One NASA" concept, MU-SPIN is working with NASA Headquarters and the University of Texas at El Paso to bring together NASA's existing stand-alone distance learning programs. Centralizing these already successful resources will provide a stronger, more cohesive program for the education community and will better utilize NASA resources.
http://mu-spin.gsfc.nasa.gov

NCCS Computational Scientist Presented IS&T Colloquium

On February 12, 2003, the Information Science and Technology (IS&T) Colloquium featured Phillip Bording of CSC, Lead Software Engineer at the NCCS. Bording presented "Application Specific Parallel Computing Versus Microprocessors and a Big Switch." He theorized that memory-related performance limitations in parallel computing can be solved by designing machines for a specific application rather than for general use. He cited his own development of a computing system for seismic modeling and reverse time imaging.

Since 2000, the IS&T Colloquium has brought leaders in the IS&T field to GSFC for presentations on cutting-edge IS&T topics. The colloquium committee is chaired by John Schnase and includes John Dorband and Jacqueline Le Moigne. The series is hosted by GSFC's Office of the Assistant Director for Information Sciences and Chief Information Officer. **http://ISandTcolloq.gsfc.nasa.gov**

Earth Alert System Press Coverage

The March 10, 2003, edition of Information Week published the article "Emergency System On Parade: Maryland Emergency Management Agency Tests GPS System for Use by First-responders," by Eric Chabrow. The article described the successful testing of the Earth Alert System (EAS) by the Maryland Emergency Management Agency during the 2003 Maryland Gubernatorial Inauguration events. EAS project manager Fred Schamann, of the ESDCD's Science Communications Technology Branch, was quoted in the article.
http://www.informationweek.com/story/IWK20030306S0018

ESDCD Updates

Instrument Incubator Program Award

Richard Lyon, an optical scientist at the ESDCD, and Jay Herman, of GSFC's Atmospheric Chemistry and Dynamics Branch, won an Instrument Incubator proposal to build a Solar Viewing Interferometer Prototype (SVIP) instrument. SVIP is being developed as a precursor to the Earth Atmosphere Solar-Occultation Imager (EASI), a mission concept studied at GSFC.

EASI would see the Earth's limb (the circular outer edge of the planet) vignetted against the Sun to obtain an entire 3–D hyperspectral image of Earth's atmosphere every 24 hours at 2 km resolution. This will enable scientists to study atmospheric water vapor, methane, carbon dioxide, molecular oxygen, ozone, and nitrous oxide lines in the 1–4 micron spectral bands for the entire Earth. SVIP will use Lyon's techniques for wavefront sensing and optical control to demonstrate the techniques that EASI would require.

The Instrument Incubator Program of NASA's Earth Science Technology Office (ESTO) awarded approximately $2.3M for this work, with GSFC's Independent Research and Development Program awarding an additional $150K. Development at GSFC will be completed in 3 years.
http://esto.nasa.gov/programs/iip/

Image Analysis Methods Award

"A Reconfigurable Computing Environment for On-Board Data Reduction and Cloud Detection" was awarded more than $1.01M in funding over 3 years. Jacqueline Le Moigne, a senior computer scientist at the ESDCD, is the Principal Investigator for this work. This project will test the feasibility of performing automatic processing of remote sensing data onboard spacecraft using a Reconfigurable Data Path Processor (RDPP). The RDPP is special reconfigurable hardware designed to withstand the unique environment of outer space. Performing data reduction onboard spacecraft would reduce the bandwidth needed to transmit data between the spacecraft and ground stations, thereby reducing costs.

Funding for this research was awarded by the Advanced Information Systems Technology (AIST) Program of NASA's Office of Earth Science. GSFC

Co-Investigators include Pen-Shu Yeh of the Microelectronics and Signal Processing Branch; Joanna Joiner of the Data Assimilation Office; and Richard Irish, SSAI, of the Biospheric Sciences Branch. The University of Idaho and George Washington University are also participating.

As an expert in image analysis, Le Moigne has made many contributions to this field of research. She recently led the development of a new method of dimension reduction, which filters out information that is redundant for a specific application. When tested on a Beowulf computing cluster, the new method's performance was validated as more computationally efficient and more accurate than Principal Component Analysis, a common technique for dimension reduction. The results of this work were presented in the paper "Multiple Sensor Image Registration, Image Fusion and Dimension Reduction of Earth Science Imagery" at the 5th International Conference on Information Fusion in July 2002.

For information or questions contact:
esdcdnews@webserv.gsfc.nasa.gov